

ELEMENTARY PROOF OF LOGARITHMIC SOBOLEV INEQUALITIES FOR GAUSSIAN CONVOLUTIONS ON \mathbb{R}

DAVID ZIMMERMANN

ABSTRACT. In a 2013 paper, the author showed that the convolution of a compactly supported measure on the real line with a Gaussian measure satisfies a logarithmic Sobolev inequality (LSI). In a 2014 paper, the author gave bounds for the optimal constants in these LSIs. In this paper, we give a simpler, elementary proof of this result.

1. INTRODUCTION

A probability measure μ on \mathbb{R}^n is said to satisfy a logarithmic Sobolev inequality (LSI) with constant $c \in \mathbb{R}$ if

$$\text{Ent}_\mu(f^2) \leq c \mathcal{E}(f, f)$$

for all locally Lipschitz functions $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$, where Ent_μ , called the entropy functional, is defined as

$$\text{Ent}_\mu(f) := \int f \log \frac{f}{\int f d\mu} d\mu$$

and $\mathcal{E}(f, f)$, the energy of f , is defined as

$$\mathcal{E}(f, f) := \int |\nabla f|^2 d\mu,$$

with $|\nabla f|$ defined as

$$|\nabla f|(x) := \limsup_{y \rightarrow x} \frac{|f(x) - f(y)|}{|x - y|}$$

so that $|\nabla f|$ is defined everywhere and coincides with the usual notion of gradient where f is differentiable. The smallest c for which a LSI with constant c holds is called the optimal log-Sobolev constant for μ .

LSIs are a useful tool that have been applied in various areas of mathematics, cf. [1, 2, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20]. In [21], the present author showed that the convolution of a compactly supported measure on \mathbb{R} with a Gaussian measure satisfies a LSI, and an application of this fact to random matrix theory was given. In [22, Thms. 2,3], bounds for the optimal constants in these LSIs were given, and the results were extended to \mathbb{R}^n . Those results are stated as Theorems 1 and 2 below. (See [17] for statements about LSIs for convolutions with more general measures).

Theorem 1. *Let μ be a probability measure on \mathbb{R} whose support is contained in an interval of length $2R$, and let γ_δ be the centered Gaussian of variance $\delta > 0$, i.e., $d\gamma_\delta(t) = (2\pi\delta)^{-1/2} \exp(-\frac{t^2}{2\delta})dt$. Then for some absolute constants K_i , the optimal log-Sobolev constant $c(\delta)$ for $\mu * \gamma_\delta$ satisfies*

$$c(\delta) \leq K_1 \frac{\delta^{3/2} R}{4R^2 + \delta} \exp\left(\frac{2R^2}{\delta}\right) + K_2 (\sqrt{\delta} + 2R)^2.$$

In particular, if $\delta \leq R^2$, then

$$c(\delta) \leq K_3 \frac{\delta^{3/2}}{R} \exp\left(\frac{2R^2}{\delta}\right).$$

The K_i can be taken in the above inequalities to be $K_1 = 6905$, $K_2 = 4989$, $K_3 = 7803$.

Theorem 2. *Let μ be a probability measure on \mathbb{R}^n whose support is contained in a ball of radius R , and let γ_δ be the centered Gaussian of variance δ with $0 < \delta \leq R^2$, i.e., $d\gamma_\delta(x) = (2\pi\delta)^{-n/2} \exp(-\frac{|x|^2}{2\delta})dx$. Then for some absolute constant K , the optimal log-Sobolev constant $c(\delta)$ for $\mu * \gamma_\delta$ satisfies*

$$c(\delta) \leq K R^2 \exp\left(20n + \frac{5R^2}{\delta}\right).$$

K can be taken above to be 289.

Theorem 1 was proved in [22] using the following theorem due to Bobkov and Götze [3, p.25, Thm 5.3]:

Theorem 3 (Bobkov, Götze). *Let μ be a Borel probability measure on \mathbb{R} with distribution function $F(x) = \mu((-\infty, x])$. Let p be the density of the absolutely continuous part of μ with respect to Lebesgue measure, and let m be a median of μ . Let*

$$D_0 = \sup_{x < m} \left(F(x) \cdot \log \frac{1}{F(x)} \cdot \int_x^m \frac{1}{p(t)} dt \right),$$

$$D_1 = \sup_{x > m} \left((1 - F(x)) \cdot \log \frac{1}{1 - F(x)} \cdot \int_m^x \frac{1}{p(t)} dt \right),$$

defining D_0 and D_1 to be zero if $\mu((-\infty, m)) = 0$ or $\mu((m, \infty)) = 0$, respectively, and using the convention $0 \cdot \infty = 0$. Then the optimal log Sobolev constant c for μ satisfies $\frac{1}{150}(D_0 + D_1) \leq c \leq 468(D_0 + D_1)$.

Theorem 2 was proved in [22] using the following theorem due to Cattiaux, Guillin, and Wu [6, Thm. 1.2]:

Theorem 4 (Cattiaux, Guillin, Wu). *Let μ be a probability measure on \mathbb{R}^n with $d\mu(x) = e^{-V(x)}dx$ for some $V \in C^2(\mathbb{R}^n)$. Suppose the following:*

- (1) *There exists a constant $K \leq 0$ such that $\text{Hess}(V) \geq KI$.*
- (2) *There exists a $W \in C^2(\mathbb{R}^n)$ with $W \geq 1$ and constants $b, c > 0$ such that*

$$\Delta W(x) - \langle \nabla V, \nabla W \rangle(x) \leq (b - c|x|^2)W(x)$$

for all $x \in \mathbb{R}^n$.

Then μ satisfies a LSI.

The goal of the present paper is to provide an elementary proof of Theorem 1. The result proved is the following:

Theorem 5. *Let μ be a probability measure on \mathbb{R} whose support is contained in an interval of length $2R$, and let γ_δ be the centered Gaussian of variance $\delta > 0$, i.e., $d\gamma_\delta(t) = (2\pi\delta)^{-1/2} \exp(-\frac{t^2}{2\delta})dt$. Then the optimal log-Sobolev constant $c(\delta)$ for $\mu * \gamma_\delta$ satisfies*

$$c(\delta) \leq \max \left(2\delta \exp \left(\frac{4R^2}{\delta} + \frac{4R}{\sqrt{\delta}} + \frac{1}{4} \right), 2\delta \exp \left(\frac{24R^2}{\delta} \right) \right).$$

In particular, if $\delta \leq 16R^2$, we have

$$c(\delta) \leq 2\delta \exp \left(\frac{24R^2}{\delta} \right).$$

The bound in Theorem 5 is worse than the bound in Theorem 1 for small δ , but still has an order of magnitude that is exponential in R^2/δ . (It is shown in [22, Example 21] that one cannot do better than exponential in R^2/δ for small δ .)

2. PROOF OF THEOREM 5

The proof of Theorem 5 is based on two facts: first, the Gaussian measure γ_1 of unit variance satisfies a LSI with constant 2. Second, Lipschitz functions preserve LSIs. We give a precise statement of this second fact below.

Proposition 6. *Let μ be a measure on \mathbb{R} that satisfies a LSI with constant c , and let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be Lipschitz. Then the push-forward measure $T_*\mu$ also satisfies a LSI with constant $c\|T\|_{\text{Lip}}^2$.*

Proof. Let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be locally Lipschitz. Then $g \circ T$ is locally Lipschitz, so by the LSI for μ ,

$$(1) \quad \int (g \circ T)^2 \log \frac{(g \circ T)^2}{\int (g \circ T)^2 d\mu} d\mu \leq c \int |\nabla(g \circ T)|^2 d\mu.$$

But since T is Lipschitz,

$$|\nabla(g \circ T)| \leq (|\nabla g| \circ T)\|T\|_{\text{Lip}}.$$

So by a change of variables, (1) simply becomes

$$\int g^2 \log \frac{g^2}{\int g^2 dT_*\mu} dT_*\mu \leq c\|T\|_{\text{Lip}}^2 \int |\nabla g|^2 dT_*\mu.$$

as desired. □

We now prove Theorem 5.

Proof of Theorem 5. In light of Proposition 6, we will establish the theorem by showing that $\mu * \gamma_\delta$ is the push-forward of γ_1 under a Lipschitz map. By translation invariance of LSI, we can assume that $\text{supp}(\mu) \subseteq [-R, R]$. We will also first assume that $\delta = 1$ (the general case will be handled at the end of the proof by a scaling argument).

Let F and G be the cumulative distribution functions of γ_1 and $\mu * \gamma_1$, i.e.,

$$F(x) = \int_{-\infty}^x p(t) dt, \quad G(x) = \int_{-\infty}^x q(t) dt,$$

where

$$p(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \quad \text{and} \quad q(t) = \int_{-R}^R p(t-s) d\mu(s).$$

Notice that q is smooth and strictly positive, so that $G^{-1} \circ F$ is well-defined and smooth. It is readily seen that $(G^{-1} \circ F)_*(\gamma_1) = \mu * \gamma_1$, so to establish the theorem we simply need to bound the derivative of $G^{-1} \circ F$.

Now

$$(G^{-1} \circ F)'(x) = \frac{1}{G'((G^{-1} \circ F)(x))} \cdot F'(x) = \frac{p(x)}{q((G^{-1} \circ F)(x))}.$$

We will bound the above derivative in cases – when $x \geq 2R$, when $-2R \leq x \leq 2R$, and when $x \leq -2R$.

We first consider the case $x \geq 2R$. Define

$$\Lambda(x) = \int_{-R}^R e^{xs} d\mu(s), \quad K(x) = \frac{\log \Lambda(x) + R}{x}.$$

Note Λ and K are smooth for $x \neq 0$.

Lemma 7. *For $x \geq 2R$,*

$$\exp\left(-2R^2 - 2R - \frac{1}{8}\right) p(x) \leq q(x + K(x)) \leq e^{-R} p(x).$$

Proof. By definition of q, p, Λ , and K ,

$$\begin{aligned} q(x + K(x)) &= \int_{-R}^R p(x + K(x) - s) d\mu(s) = p(x) \cdot e^{-xK(x)} \int_{-R}^R \exp\left(-\frac{(K(x) - s)^2}{2}\right) \cdot e^{xs} d\mu(s) \\ &= \frac{e^{-R} p(x)}{\Lambda(x)} \int_{-R}^R \exp\left(-\frac{(K(x) - s)^2}{2}\right) \cdot e^{xs} d\mu(s) \\ &\leq \frac{e^{-R} p(x)}{\Lambda(x)} \int_{-R}^R e^{xs} d\mu(s) \\ &= e^{-R} p(x). \end{aligned}$$

To get the other inequality, first note that $e^{-Rx} \leq \Lambda(x) \leq e^{Rx}$. (These are just the maximum and minimum values in the integrand defining Λ .) This implies that $-R + R/x \leq K(x) \leq R + R/x$, so for $-R \leq s \leq R$ and $x \geq 2R$, we have

$$-2R - \frac{R}{x} \leq -2R + \frac{R}{x} \leq K(x) - s \leq 2R + \frac{R}{x}$$

so that

$$\exp\left(-\frac{(K(x) - s)^2}{2}\right) \geq \exp\left(-\frac{(2R + R/x)^2}{2}\right) \geq \exp\left(-\frac{(2R + R/(2R))^2}{2}\right) = \exp\left(-2R^2 - R - \frac{1}{8}\right).$$

Therefore

$$q(x + K(x)) = \frac{e^{-R} p(x)}{\Lambda(x)} \int_{-R}^R \exp\left(-\frac{(K(x) - s)^2}{2}\right) \cdot e^{xs} d\mu(s) \geq \exp\left(-2R^2 - 2R - \frac{1}{8}\right) p(x).$$

□

Lemma 8. $K'(x) \leq R$ for $x \geq 2R$.

Proof. Recall that $e^{-Rx} \leq \Lambda(x)$. (Again, e^{-Rx} is the minimum value in the integrand defining Λ). We therefore have

$$\begin{aligned} K'(x) &= \frac{\Lambda'(x)}{x\Lambda(x)} - \frac{\log \Lambda(x)}{x^2} - \frac{R}{x^2} = \frac{\int_{-R}^R s e^{sx} d\mu(s)}{x\Lambda(x)} - \frac{\log \Lambda(x)}{x^2} - \frac{R}{x^2} \\ &\leq \frac{R \int_{-R}^R e^{sx} d\mu(s)}{x\Lambda(x)} + \frac{Rx}{x^2} - \frac{R}{x^2} \\ &= \frac{2R}{x} - \frac{R}{x^2}. \end{aligned}$$

By elementary calculus, the above has a maximum value of R . □

Lemma 9. For $x \geq 2R$,

$$x - R \leq (G^{-1} \circ F)(x) \leq x + K(x).$$

Proof. Since G and G^{-1} are increasing, the lemma is equivalent to

$$G(x - R) \leq F(x) \leq G(x + K(x)).$$

The first inequality follows from the definition of G and the Fubini-Tonelli Theorem:

$$\begin{aligned} G(x - R) &= \int_{-\infty}^{x-R} q(t) dt = \int_{-\infty}^x \int_{-R}^R p(t - s) d\mu(s) dt = \int_{-R}^R \int_{-\infty}^{x-R} p(t - s) dt d\mu(s) \\ &= \int_{-R}^R \int_{-\infty}^{x-R+s} p(u) du d\mu(s) \\ &\quad \text{where } u = t - s \\ &\leq \int_{-R}^R \int_{-\infty}^x p(u) dt d\mu(s) \\ &= F(x). \end{aligned}$$

To establish the other inequality, we use Lemmas 7 and 8:

$$\begin{aligned} 1 - G(x + K(x)) &= \int_{x+K(x)}^{\infty} q(t) dt = \int_x^{\infty} q(u + K(u))(1 + K'(u)) du \\ &\quad \text{where } t = u + K(u) \\ &\leq \int_x^{\infty} p(u) e^{-R}(1 + R) du \\ &\quad \text{by Lemmas 7 and 8} \\ &\leq \int_x^{\infty} p(u) du \\ &\quad \text{since } e^R \geq 1 + R \\ &= 1 - F(x), \end{aligned}$$

so that $F(x) \leq G(x + K(x))$, as desired. □

We are almost ready to bound $(G^{-1} \circ F)'(x)$ for $x \geq 2R$. The last observation to make is that q is decreasing on $[R, \infty)$ since

$$q'(t) = \int_{-R}^R p'(t - s) d\mu(s) = \int_{-R}^R -(t - s)p(t - s) d\mu(s) \leq 0 \quad \text{for } t \geq R.$$

So for $x \geq 2R$ we have, by lemma 9,

$$q((G^{-1} \circ F)(x)) \geq q(x + K(x)).$$

Combining this with Lemma 7, we get

$$(G^{-1} \circ F)'(x) = \frac{p(x)}{q((G^{-1} \circ F)(x))} \leq \frac{p(x)}{q(x + K(x))} \leq \exp\left(2R^2 + 2R + \frac{1}{8}\right)$$

for $x \geq 2R$.

In the case where $-2R \leq x \leq 2R$, first note that for all x ,

$$x - R \leq (G^{-1} \circ F)(x) \leq x + R;$$

the first inequality above was done in Lemma 9, and the second inequality is proven in the same way. So

$$\sup_{-2R \leq x \leq 2R} (G^{-1} \circ F)'(x) = \sup_{-2R \leq x \leq 2R} \frac{p(x)}{q((G^{-1} \circ F)(x))} \leq \sup_{\substack{-2R \leq x \leq 2R \\ -R \leq y \leq R}} \frac{p(x)}{q(x+y)} = \left(\inf_{\substack{-2R \leq x \leq 2R \\ -R \leq y \leq R}} \frac{q(x+y)}{p(x)} \right)^{-1}.$$

For convenience, let $S = \{(x, y) : -2R \leq x \leq 2R, -R \leq y \leq R\}$. Now

$$\inf_{(x,y) \in S} \frac{q(x+y)}{p(x)} = \inf_{(x,y) \in S} \frac{1}{p(x)} \int_{-R}^R p(x+y-s) d\mu(s).$$

Since p has no local minima, the minimum value of the above integrand occurs at either $s = R$ or $s = -R$. Without loss of generality, we assume the minimum is achieved at $s = R$ (otherwise, we can replace (x, y) with $(-x, -y)$ by symmetry of S and p). So

$$\inf_{(x,y) \in S} \frac{q(x+y)}{p(x)} \geq \inf_{(x,y) \in S} \frac{1}{p(x)} \cdot p(x+y+R).$$

Elementary calculus shows that the above infimum is equal to e^{-12R^2} (achieved at $x = 2R, y = R$). Therefore

$$\sup_{-2R \leq x \leq 2R} (G^{-1} \circ F)'(x) = \left(\inf_{(x,y) \in S} \frac{q(x+y)}{p(x)} \right)^{-1} \leq e^{12R^2}.$$

The case $x \leq -2R$ is dealt with in the same way as the case $x \geq 2R$, the analagous statements being:

$$\exp\left(-2R^2 - 2R - \frac{1}{8}\right) p(x) \leq q(x+K(x)) \leq e^{-R} p(x),$$

$$K'(x) \leq R,$$

$$x + K(x) \leq (G^{-1} \circ F)(x) \leq x + R,$$

and q is increasing for $x \leq -2R$. The upper bound for $(G^{-1} \circ F)'(x)$ obtained in this case is the same as the one in the case $x \geq 2R$.

We therefore have

$$\|G^{-1} \circ F\|_{\text{Lip}} \leq \max\left(\exp\left(2R^2 + 2R + \frac{1}{8}\right), e^{12R^2}\right)$$

So by Proposition 6, $\mu * \gamma_1$ satisfies a LSI with constant $c(1)$ satisfying

$$c(1) \leq 2\|G^{-1} \circ F\|_{\text{Lip}}^2 \leq \max\left(2\exp\left(4R^2 + 4R + \frac{1}{4}\right), 2e^{24R^2}\right).$$

This proves the theorem for the case $\delta = 1$.

To establish the theorem for a general $\delta > 0$, first observe that

$$\mu * \gamma_\delta = (h_{\sqrt{\delta}})_* \left(((h_{1/\sqrt{\delta}})_* \mu) * \gamma_1 \right),$$

where h_λ denotes the scaling map with factor λ , i.e., $h_\lambda(x) = \lambda x$. Now $(h_{1/\sqrt{\delta}})_* \mu$ is supported in $[-R/\sqrt{\delta}, R/\sqrt{\delta}]$, so by the case $\delta = 1$ just proven, $((h_{1/\sqrt{\delta}})_* \mu) * \gamma_1$ satisfies a LSI with constant

$$\max\left(2\exp\left(4(R/\sqrt{\delta})^2 + 4(R/\sqrt{\delta}) + \frac{1}{4}\right), 2e^{24(R/\sqrt{\delta})^2}\right).$$

Finally, since $\|h_{\sqrt{\delta}}\|_{\text{Lip}}^2 = \delta$, we have by Proposition 6,

$$c(\delta) \leq \max\left(2\delta \exp\left(\frac{4R^2}{\delta} + \frac{4R}{\sqrt{\delta}} + \frac{1}{4}\right), 2\delta \exp\left(\frac{24R^2}{\delta}\right)\right).$$

In particular, when $\delta \leq 16R^2$, we have

$$2\delta \exp\left(\frac{4R^2}{\delta} + \frac{4R}{\sqrt{\delta}} + \frac{1}{4}\right) \leq 2\delta \exp\left(\frac{24R^2}{\delta}\right)$$

so the above bound on $c(\delta)$ simplifies to

$$c(\delta) \leq 2\delta \exp\left(\frac{24R^2}{\delta}\right).$$

□

ACKNOWLEDGEMENTS

The author would like to thank his Ph.D. advisor, Todd Kemp, for his valuable insights and discussions regarding this topic.

REFERENCES

- [1] Bakry, D.: *L'hypercontractivité et son utilisation en thorie des semigroupes*. Lectures on probability theory (Saint-Flour, 1992), 1–114, Lecture Notes in Math., 1581, Springer, Berlin, 1994.
- [2] Bakry, D.: *On Sobolev and logarithmic Sobolev inequalities for Markov semigroups*. New trends in stochastic analysis (Charingworth, 1994), 43–75, World Sci. Publ., River Edge, NJ, 1997.
- [3] Bobkov, S. G.; Götze, F.: *Exponential Integrability and Transportation Cost Related to Logarithmic Sobolev Inequalities*, J. Funct. Anal. **163** (1999), 1–28.
- [4] Bobkov, S.; Houdré, C.: *Some connections between isoperimetric and Sobolev-type inequalities*. Mem. Amer. Math. Soc. **129** no. 616 (1997)
- [5] Bobkov, S.; Tetali, P.: *Modied logarithmic Sobolev inequalities in discrete settings*. J. Theoret. Probab. **19**, no. 2, 289–336 (2006)
- [6] Cattiaux, P.; Guillin, A.; Wu, L.: *A note on Talagrand's transportation inequality and logarithmic Sobolev inequality*. Probab. Theory Relat. Fields **148**, 285–304 (2010)
- [7] Davies, E. B.: *Explicit constants for Gaussian upper bounds on heat kernels*. Amer. J. Math. **109**, no. 2, 319–333 (1987)
- [8] Davies, E. B.: *Heat kernels and spectral theory*. Cambridge Tracts in Mathematics, 92. Cambridge University Press, 1990.
- [9] Davies, E. B., Simon, B.: *Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians*. J. Funct. Anal. **59**, 335–395 (1984)
- [10] Diaconis, P., Saloff-Coste, L.: *Logarithmic Sobolev inequalities for finite Markov chains*. Ann. Appl. Probab. **6**, 695–750 (1996)
- [11] Gross, L.; Rothaus, O.: *Herbst inequalities for supercontractive semigroups*. J. Math. Kyoto Univ. **38**, no. 2, 295–318 (1998)
- [12] Guionnet, A.; Zegarlinski, B.: *Lectures on logarithmic Sobolev inequalities*. Séminaire de Probabilités, XXXVI, 1134, Lecture Notes in Math., 1801, Springer, Berlin, 2003.
- [13] Ledoux, M.: *Isoperimetry and Gaussian analysis. Lectures on probability theory and statistics*. 165–294, Lecture Notes in Math 1648, Springer, Berlin, 1996
- [14] Ledoux, M.: *The concentration of measure phenomenon*. Mathematical Surveys and Monographs, 89. American Mathematical Society, Providence, RI, 2001.
- [15] Ledoux, M.: *A remark on hypercontractivity and tail inequalities for the largest eigenvalues of random matrices*. Séminaire de Probabilités XXXVII, 360–369, Lecture Notes in Math., 1832, Springer, Berlin, 2003.
- [16] Villani, C.: *Topics in optimal transportation*. Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, 2003.
- [17] Wang, F.-Y.; Wang, J.: *Functional inequalities for convolution probability measures*. arXiv:1308.1713
- [18] Yau, H.T.: *Logarithmic Sobolev inequality for the lattice gases with mixing conditions*. Commun. Math. Phys. **181**, 367–408 (1996)
- [19] Yau, H.T.: *Log-Sobolev inequality for generalized simple exclusion processes*. Probab. Theory Related Fields **109**, 507–538 (1997)
- [20] Zegarlinski, B.: *Dobrushin uniqueness theorem and logarithmic Sobolev inequalities*. J. Funct. Anal. **105**, 77–111 (1992)
- [21] Zimmermann, D.: *Logarithmic Sobolev inequalities for mollified compactly supported measures*. J. Funct. Anal. **265**, 1064–1083 (2013)
- [22] Zimmermann, D.: *Bounds for logarithmic Sobolev constants for Gaussian convolutions of compactly supported measures*. arXiv:1405.2581

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SAN DIEGO 92093
E-mail address: dszimmer@math.ucsd.edu